

**ANALISIS SENTIMEN PADA *REVIEW* PRODUK KOSMETIK
MENGUNAKAN METODE *MULTINOMIAL NAÏVE BAYES*
DILENGKAPI SELEKSI FITUR *INFORMATION GAIN***

SKRIPSI

**Diajukan untuk Memenuhi Salah Satu Syarat Mencapai Gelar Strata Satu
Program Studi Informatika**



Disusun Oleh:

DIANI NUR PRIHANTARI FAUZI

NIM M0512013

**PROGRAM STUDI INFORMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS SEBELAS MARET
SURAKARTA**

2017

SKRIPSI

**ANALISIS SENTIMEN PADA *REVIEW* PRODUK KOSMETIK
MENGUNAKAN METODE *MULTINOMIAL NAÏVE BAYES*
DILENGKAPI SELEKSI FITUR *INFORMATION GAIN***

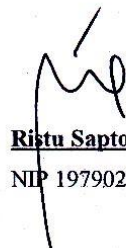
Disusun oleh:

DIANI NUR PRIHANTARI FAUZI

NIM M0512013

telah disetujui untuk dipertahankan di hadapan Dewan Penguji
pada tanggal: 10 April 2017

Pembimbing I,



Ristu Saptono S.Si., M.T.

NIP 197902102002121001

Pembimbing II,



Sari Widya Sihwi, S.Kom., M.T.I.

NIP 198304122009122003

SKRIPSI

**ANALISIS SENTIMEN PADA *REVIEW* PRODUK KOSMETIK
MENGUNAKAN METODE *MULTINOMIAL NAÏVE BAYES*
DILENGKAPI SELEKSI FITUR *INFORMATION GAIN***

Disusun oleh:


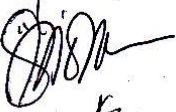


DIANI NUR PRIHANTARI FAUZI

NIM M0512013

telah dipertahankan di hadapan Dewan Penguji

pada tanggal: 3 Mei 2017

Susunan Dewan Penguji

1. **Ristu Saptono S.Si., M.T.** ()
NIP. 19790210 200212 1 001
2. **Sari Widya Sihwi, S.Kom., M.T.I.** ()
NIP. 19830412 200912 2 003
3. **Esti Survani, S.Si., M.Kom.** ()
NIP. 19761129 200812 2 001
4. **Afrizal Doewes, S.Kom., M.Sc.** ()
NIP. 19850831 200212 1 004

Disahkan oleh

Kepala Program Studi Informatika,


Drs. Bambang Harjito, M.App.Sc., Ph.D.

NIP. 19621130 199103 1 002

MOTTO

“Fabiayyi aalaa i rabbikumaa tukadzdzibaan”

Maka nikmat Tuhanmu yang manakah yang kamu dustakan?

(QS 55:13)

“Fa inna ma’al ‘usri yusra. Inna ma’al ‘usri yusra”

Maka sesungguhnya bersama kesulitan ada kemudahan. Sesungguhnya bersama kesulitan ada kemudahan.

(QS 94:6-7)

“La tahzan innallaha ma’ana”

Jangan engkau bersedih, sesungguhnya Allah bersama kita.

(QS 9:40)

PERSEMBAHAN

Karya ini penulis persembahkan kepada :

Mama, Papa, dan adik-adikku,

Teman-teman Informatika 2012,

Gannon Bridung Errinta Putra.

KATA PENGANTAR

Puji syukur kehadiran Allah Subhanahu Wa Ta'ala atas limpahan berkah dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul Analisis Sentimen pada *Review* Produk Kosmetik Menggunakan Metode *Multinomial Naïve Bayes* Dilengkapi Seleksi Fitur *Information Gain*.

Penulis menyadari bahwa masih banyak kekurangan dalam skripsi ini. Begitu banyak bantuan, dukungan, serta bimbingan yang penulis dapatkan. Maka penulis mengucapkan terimakasih kepada :

1. Mama, Papa, dan adik-adik yang telah memberikan dukungan dan pendampingan selama ini.
2. Bapak Drs. Bambang Harjito, M.App.Sc., Ph.D., selaku Kepala Program Studi Informatika yang telah memberi motivasi dan dukungan selama proses perkuliahan
3. Bapak Ristu Saptono, S.Si., M.T., selaku Dosen Pembimbing I dan juga Pembimbing Akademik yang telah memberi dukungan, motivasi, masukan, dan kritik yang membangun selama proses perkuliahan dan penyusunan skripsi
4. Ibu Sari Widya Sihwi, S.Kom., M.T.I. selaku Dosen Pembimbing II yang telah memberi dukungan dan motivasi selama proses penyusunan skripsi
5. Bapak dan Ibu dosen Informatika beserta karyawan yang telah memberi pengajaran dan bantuan selama masa perkuliahan dan penyusunan skripsi
6. Amel, Zainal, Ardi, dan Damas yang membantu penulis dalam proses penyusunan skripsi
7. Teman-teman yang memberikan dukungan dan motivasi kepada penulis
8. Mas Gannon yang selalu memberikan dukungan, motivasi, dan semangat

Penulis berharap semoga skripsi ini bermanfaat bagi penulis dan semua pihak yang berkepentingan.

Surakarta, _____ 2017

Penulis

ANALISIS SENTIMEN PADA *REVIEW* PRODUK KOSMETIK MENGUNAKAN METODE *MULTINOMIAL NAÏVE BAYES* DILENGKAPI SELEKSI FITUR *INFORMATION GAIN*

DIANI NUR PRIHANTARI FAUZI

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Sebelas Maret

ABSTRAK

Di Indonesia, terdapat media bernama *femaledaily.com* yang memiliki kolom khusus untuk *review* produk kosmetik. *Review* produk dapat membantu konsumen menyimpulkan kualitas produk. Beberapa cara untuk menyimpulkan kualitas produk adalah dengan mengklasifikasikan *review* dan menghitung polaritas sentimen yang menjadi bagian dari tujuan penelitian ini. Penelitian ini bertujuan untuk mengetahui akurasi klasifikasi menggunakan metode *Multinomial Naïve Bayes* (MNB) dilengkapi seleksi fitur *Information Gain* (IG) dan menganalisa polaritas sentimen masyarakat terhadap tujuh produk kosmetik populer pada *Female Daily*.

Klasifikasi menggunakan MNB dan IG menghasilkan akurasi yang sebanding dengan akurasi klasifikasi hanya menggunakan MNB. Waktu eksekusi MNB-IG menurun separuh dari waktu eksekusi MNB. Hal ini terjadi karena IG menyebabkan berkurangnya dimensi data sehingga berpengaruh pada berkurangnya waktu yang diperlukan untuk klasifikasi. Hasil akurasi MNB-IG adalah 80,42% dan klasifikasi hanya menggunakan MNB adalah 82,17% dengan waktu eksekusi sekitar 4,7 detik untuk MNB-IG dan 9,51 detik untuk MNB. Polaritas sentimen masyarakat terhadap tujuh produk kosmetik memperlihatkan bahwa polaritas tertinggi dimiliki oleh Estee Lauder dengan polaritas positif sebesar 94,37%. Perangkingan polaritas ternyata memberikan hasil yang sebanding dengan perangkingan *user rating* dibuktikan dengan urutan polaritas dan urutan *rating* yang ternyata sama. Oleh karena itu, polaritas sentimen sebenarnya dapat diajukan kepada FemaleDaily untuk dijadikan alternatif yang lebih objektif dalam perankingan.

Kata kunci: Analisis Sentimen, *Information Gain*, *Multinomial Naïve Bayes*, *Produk Kosmetik*

SENTIMENT ANALYSIS OF COSMETIC PRODUCTS REVIEW USING MULTINOMIAL NAÏVE BAYES WITH INFORMATION GAIN FOR FEATURE SELECTION

DIANI NUR PRIHANTARI FAUZI

*Informatics Department. Faculty of Mathematics and Natural Sciences.
Sebelas Maret University*

ABSTRACT

In Indonesia, there is a website containing special column for cosmetic product reviews named femaledaily.com. A product review can be used to determine the sentiment polarity of the product itself. This study aims to determine the accuracy of Multinomial Naïve Bayes (MNB) classifier with Information Gain (IG) for feature selection and to analyze the public sentiment towards seven popular cosmetic products on Female Daily.

The classification using MNB-IG resulted comparable accuration with the classification only using MNB. The execution time of the classification using MNB-IG decreased a half than the execution time of the classification only using MNB. This happened because Information Gain caused a reduction of data dimensions so that the time needed for classification decreased. The classification accuracy of MNB-IG was 80,42% and the accuracy of MNB was 82,17%. The execution time approximately 4.7 seconds for MNB-IG and 9.51 seconds for MNB. The public sentiment towards seven popular cosmetic products showed that the highest positive polarity is owned by Estee Lauder with 94,37%. Surprisingly, the rank resulted by sentiment polarity was identical with the rank resulted by user rating. Therefore sentiment polarity can be offered to FemaleDaily for objective alternatives aside from rating.

Keywords: *Cosmetic Products, Information Gain, Multinomial Naïve Bayes, Sentiment Analysis.*

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN.....	iii
MOTTO	iv
PERSEMBAHAN.....	v
KATA PENGANTAR	vi
ABSTRAK.....	vii
<i>ABSTRACT</i>	viii
DAFTAR ISI.....	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL.....	xii
BAB I.....	1
PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
1.6 Sistematika Penulisan	4
BAB II.....	5
TINJAUAN PUSTAKA	5
2.1 Landasan Teori	5
BAB III	12
METODOLOGI PENELITIAN.....	12
BAB IV	17
HASIL DAN PEMBAHASAN.....	17
4.1 Deskripsi Data	17

4.2 Text Preprocessing	18
4.3 Penerapan Seleksi Fitur <i>Information Gain</i>	19
4.4 Perhitungan Multinomial Naïve Bayes.....	22
BAB V	28
PENUTUP	28
5.1 Simpulan.....	28
5.2 Saran	28
DAFTAR PUSTAKA	29
LAMPIRAN.....	31

DAFTAR GAMBAR

Gambar 3.1 Metodologi Penelitian.....	12
Gambar 3.1 Contoh Penerapan <i>Case Folding</i>	13
Gambar 3.2 Contoh Penerapan <i>Tokenization</i>	14
Gambar 3.3 Contoh Penerapan <i>Stemming</i>	14
Gambar 4.1 Contoh <i>review</i>	17

DAFTAR TABEL

Tabel 2.2 PenelitianTerkait	9
Tabel 2.3 Kerangka Pemikiran.....	11
Tabel 3.1 <i>Confusion Matrix</i>	16
Tabel 4.1 Seleksi Fitur	20
Tabel 4.2 Efektivitas Seleksi Fitur dan Klasifikasi	21
Tabel 4.3 Contoh <i>Data Training</i> dan <i>Data Testing</i> setelah Seleksi Fitur	22
Tabel 4.4 <i>Memorizing</i>	25
Tabel 4.5 <i>Data Testing</i>	25
Tabel 4.6 Persebaran dan Polaritas Data	26